

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/240,410

OIPE

DATE: 02/10/1999
TIME: 15:52:36

INPUT SET: S30581.raw

This Raw Listing contains the General
Information Section and up to the first 5 pages.

SEQUENCE LISTING

(1) General Information

(i) APPLICANT: MICHALOVICH, DAVID
HAYES, PHILIP DAVID

(ii) TITLE OF THE INVENTION: NOVEL COMPOUNDS

(iii) NUMBER OF SEQUENCES: 4

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Ratner & Prestia
(B) STREET: P.O. Box 980
(C) CITY: Valley Forge
(D) STATE: PA
(E) COUNTRY: USA
(F) ZIP: 19482

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: TO BE ASSIGNED
(B) FILING DATE: 27-JAN-1999
(C) CLASSIFICATION: UNKNOWN

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: UK APPLICATION NO. TO BE ASSIGNED
(B) FILING DATE: 20-JAN-1999

(A) APPLICATION NUMBER: EP APPLICATION NO. 98300694.1
(B) FILING DATE: 30-JAN-1998

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Prestia, Paul F
(B) REGISTRATION NUMBER: 23,031
(C) REFERENCE/DOCKET NUMBER: GP-30039

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 610-407-0700

ENTERED

RECEIVED
JUL 28 1999
TC 2700 MAIL ROOM

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/240,410DATE: 02/10/1999
TIME: 15:52:36

INPUT SET: S30581.raw

47 (B) TELEFAX: 610-407-0700
48 (C) TELEX: 846169
49
50

51 (2) INFORMATION FOR SEQ ID NO:1:
52

53 (i) SEQUENCE CHARACTERISTICS:
54 (A) LENGTH: 2010 base pairs
55 (B) TYPE: nucleic acid
56 (C) STRANDEDNESS: single
57 (D) TOPOLOGY: linear
58

59 (ii) MOLECULE TYPE: cDNA
60

61 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
62

63	ATGACGCCTC	CTCCGCCCCG	ACGTGCCGCC	CCCAGCGCAC	CGCGCGCCCC	CGTCCCTGGC	60
64	CCGCGGGCTC	GGTTGGGGCT	TCCGCTGCGG	CTGCGGCTGC	TGCTGCTGCT	CTGGGCGGCC	120
65	GCCGCCTCCG	CCCAGGGCCA	CCTAAGGAGC	GGACCCCGCA	TCTTCGCCGT	CTGGAAAGGC	180
66	CATGTAGGGC	AGGACCGGGT	GGACTTTGGC	CAGACTGAGC	CGCACACGGT	GCTTTTCCAC	240
67	GAGCCAGGCA	GCTCCTCTGT	GTGGGTGGGA	GGACGTGGCA	AGGTCTACCT	CTTTGACTTC	300
68	CCCAGGGGCA	AGAACGCATC	TGTGCGCACG	GTGAATATCG	GCTCCACAAA	GGGGTCCTGT	360
69	CTGGATAAGC	GGGACTGCGA	GAACATACATC	ACTCTCCTGG	AGAGGCGGAG	TGAGGGGCTG	420
70	CTGGCCTGTG	GCACCAACGC	CCGGCACCCC	AGCTGCTGGA	ACCTGGTGAA	TGGCACTGTG	480
71	GTGCCACTTG	GCGAGATGAG	AGGCTACGCC	CCCTTCAGCC	CGGACGAGAA	CTCCCTGGTT	540
72	CTGTTTGAAG	GGGACGAGGT	GTATTCCACC	ATCCGGAAGC	AGGAATACAA	TGGGAAGATC	600
73	CCTCGGTTCC	GCCGCATCCG	GGGCGAGAGT	GAGCTGTACA	CCAGTGATAC	TGTCATGCAG	660
74	AACCCACAGT	TCATCAAAGC	CACCATCGTG	CACCAAGACC	AGGCTTACGA	TGACAAGATC	720
75	TACTACTTCT	TCCGAGAGGA	CAATCCTGAC	AAGAATCCTG	AGGCTCCTCT	CAATGTGTCC	780
76	CGTGTGGCCC	AGTTGTGCAG	GGGGGACCAG	GGTGGGGAAA	GTTCACGTGC	AGTCTCCAAG	840
77	TGGAACACTT	TTCTGAAAGC	CATGCTGGTA	TGCAGTGATG	CTGCCACCAA	CAAGAACTTC	900
78	AACAGGCTGC	AAGACGTCTT	CCTGCTCCCT	GACCCACGCG	GCCAGTGGAG	GGACACCAGG	960
79	GTCTATGGTG	TTTTCTCCAA	CCCCTGGAAC	TACTCAGCCG	TCTGTGTGTA	TTCCCTCGGT	1020
80	GACATTGACA	AGGTCTTCCG	TACCTCCTCA	CTCAAGGGCT	ACCACTCAAG	CCTTCCCAAC	1080
81	CCGCGGCCTG	GCAAGTGCCT	CCCAGACCAG	CAGCCGATAC	CCACAGAGAC	CTTCCAGGTG	1140
82	GCTGACCGTC	ACCCAGAGGT	GGCGCAGAGG	GTGGAGCCCA	TGGGGCCTCT	GAAGACGCCA	1200
83	TTGTTCCACT	CTAAATACCA	CTACCAGAAA	GTGGCCGTCC	ACCGCATGCA	AGCCAGCCAC	1260
84	GGGGAGACCT	TTCATGTGCT	TTACCTAACT	ACAGACAGGG	GCACTATCCA	CAAGGTGGTG	1320
85	GAACCGGGGG	AGCAGGAGCA	CAGCTTCGCC	TTCAACATCA	TGGAGATCCA	GCCCTTCCGC	1380
86	CGCGCGGCTG	CCATCCAGAC	CATGTCGCTG	GATGCTGAGC	GGAGGAAGCT	GTATGTGAGC	1440
87	TCCCAGTGGG	AGGTGAGCCA	GGTGCCCCCTG	GACCTGTGTG	AGGTCTATGG	CGGGGGCTGC	1500
88	CACGGTTGCC	TCATGTCCCG	AGACCCCTAC	TGCGGCTGGG	ACCAAGGCCG	CTGCATCTCC	1560
89	ATCTACAGCT	CCGAACGGTC	AGTGCTGCAA	TCCATTAATC	CAGCCGAGCC	ACACAAGGAG	1620
90	TGTCCAACCC	CCAAACCAGA	CAAGGCCCCA	CTGCAGAAGG	TTTCCCTGGC	CCCAAACCTCT	1680
91	CGCTACTACC	TGAGCTGCCC	CATGGAATCC	CGCCACGCCA	CCTACTCATG	GCGCCACAAG	1740
92	GAGAACGTGG	AGCAGAGCTG	CGAACCTGGT	CACCAGAGCC	CCAACCTGCAT	CCTGTTCATC	1800
93	GAGAACCTCA	CGGCGCAGCA	GTACGGCCAC	TACTTCTGCG	AGGCCCAGGA	GGGCTCCTAC	1860
94	TTCCGCGAGG	CTCAGCACTG	GCAGCTGCTG	CCCGAGGACG	GCATCATGGC	CGAGCACCTG	1920
95	CTGGGTTCATG	CCTGTGCCCT	GGCCGCCTCC	CTCTGGCTGG	GGGTGCTGCC	CACACTCACT	1980
96	CTTGGCTTGC	TGGTCCACTA	GGGCCTCCCC				2010

97
98 (2) INFORMATION FOR SEQ ID NO:2:
99

RECEIVED
JUL 28 1999
TC 2700 MAIL ROOM

RAW SEQUENCE LISTING PATENT APPLICATION US/08/240,410

DATE: 02/10/1999
TIME: 15:52:36

INPUT SET: S30581.raw

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 666 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

100	Met	Thr	Pro	Pro	Pro	Pro	Gly	Arg	Ala	Ala	Pro	Ser	Ala	Pro	Arg	Ala
101	1				5					10					15	
102																
103																
104																
105																
106																
107																
108																
109																
110	Arg	Val	Pro	Gly	Pro	Pro	Ala	Arg	Leu	Gly	Leu	Pro	Leu	Arg	Leu	Arg
111				20					25				30			
112																
113	Leu	Leu	Leu	Leu	Leu	Trp	Ala	Ala	Ala	Ala	Ser	Ala	Gln	Gly	His	Leu
114				35				40					45			
115	Arg	Ser	Gly	Pro	Arg	Ile	Phe	Ala	Val	Trp	Lys	Gly	His	Val	Gly	Gln
116																
117																
118	Asp	Arg	Val	Asp	Phe	Gly	Gln	Thr	Glu	Pro	His	Thr	Val	Leu	Phe	His
119																
120																
121	Glu	Pro	Gly	Ser	Ser	Ser	Val	Trp	Val	Gly	Arg	Gly	Lys	Val	Tyr	
122																
123																
124	Leu	Phe	Asp	Phe	Pro	Glu	Gly	Lys	Asn	Ala	Ser	Val	Arg	Thr	Val	Asn
125																
126	Ile	Gly	Ser	Thr	Lys	Gly	Ser	Cys	Leu	Asp	Lys	Arg	Asp	Cys	Glu	Asn
127																
128																
129	Tyr	Ile	Thr	Leu	Leu	Glu	Arg	Arg	Ser	Glu	Gly	Leu	Leu	Ala	Cys	Gly
130																
131																
132	Thr	Asn	Ala	Arg	His	Pro	Ser	Cys	Trp	Asn	Leu	Val	Asn	Gly	Thr	Val
133																
134																
135	Val	Pro	Leu	Gly	Glu	Met	Arg	Gly	Tyr	Ala	Pro	Phe	Ser	Pro	Asp	Glu
136																
137																
138	Asn	Ser	Leu	Val	Leu	Phe	Glu	Gly	Asp	Glu	Val	Tyr	Ser	Thr	Ile	Arg
139																
140																
141	Lys	Gln	Glu	Tyr	Asn	Gly	Lys	Ile	Pro	Arg	Phe	Arg	Arg	Ile	Arg	Gly
142																
143																
144																
145	Glu	Ser	Glu	Leu	Tyr	Thr	Ser	Asp	Thr	Val	Met	Gln	Asn	Pro	Gln	Phe
146																
147																
148	Ile	Lys	Ala	Thr	Ile	Val	His	Gln	Asp	Gln	Ala	Tyr	Asp	Asp	Lys	Ile
149																
150																
151	Tyr	Tyr	Phe	Phe	Arg	Glu	Asp	Asn	Pro	Asp	Lys	Asn	Pro	Glu	Ala	Pro
152																

RECEIVED
JUL 28 1999
TC 2700 MAIL ROOM

RAW SEQUENCE LISTING PATENT APPLICATION US/08/240,410

DATE: 02/10/1999
TIME: 15:52:37

INPUT SET: S30581.raw

```

153 Tyr Ser Leu Gly Asp Ile Asp Lys Val Phe Arg Thr Ser Ser Leu Lys
154          340          345          350
155 Gly Tyr His Ser Ser Leu Pro Asn Pro Arg Pro Gly Lys Cys Leu Pro
156          355          360          365
157 Asp Gln Gln Pro Ile Pro Thr Glu Thr Phe Gln Val Ala Asp Arg His
158          370          375          380
159 Pro Glu Val Ala Gln Arg Val Glu Pro Met Gly Pro Leu Lys Thr Pro
160          385          390          395          400
161 Leu Phe His Ser Lys Tyr His Tyr Gln Lys Val Ala Val His Arg Met
162          405          410          415
163 Gln Ala Ser His Gly Glu Thr Phe His Val Leu Tyr Leu Thr Thr Asp
164          420          425          430
165 Arg Gly Thr Ile His Lys Val Val Glu Pro Gly Glu Gln Glu His Ser
166          435          440          445
167 Phe Ala Phe Asn Ile Met Glu Ile Gln Pro Phe Arg Arg Ala Ala Ala
168          450          455          460
169
170 Ile Gln Thr Met Ser Leu Asp Ala Glu Arg Arg Lys Leu Tyr Val Ser
171          465          470          475          480
172 Ser Gln Trp Glu Val Ser Gln Val Pro Leu Asp Leu Cys Glu Val Tyr
173          485          490          495
174 Gly Gly Gly Cys His Gly Cys Leu Met Ser Arg Asp Pro Tyr Cys Gly
175          500          505          510
176 Trp Asp Gln Gly Arg Cys Ile Ser Ile Tyr Ser Ser Glu Arg Ser Val
177          515          520          525
178 Leu Gln Ser Ile Asn Pro Ala Glu Pro His Lys Glu Cys Pro Asn Pro
179          530          535          540
180 Lys Pro Asp Lys Ala Pro Leu Gln Lys Val Ser Leu Ala Pro Asn Ser
181          545          550          555          560
182 Arg Tyr Tyr Leu Ser Cys Pro Met Glu Ser Arg His Ala Thr Tyr Ser
183          565          570          575
184 Trp Arg His Lys Glu Asn Val Glu Gln Ser Cys Glu Pro Gly His Gln
185          580          585          590
186 Ser Pro Asn Cys Ile Leu Phe Ile Glu Asn Leu Thr Ala Gln Gln Tyr
187          595          600          605
188 Gly His Tyr Phe Cys Glu Ala Gln Glu Gly Ser Tyr Phe Arg Glu Ala
189          610          615          620
190 Gln His Trp Gln Leu Leu Pro Glu Asp Gly Ile Met Ala Glu His Leu
191          625          630          635          640
192 Leu Gly His Ala Cys Ala Leu Ala Ala Ser Leu Trp Leu Gly Val Leu
193          645          650          655
194 Pro Thr Leu Thr Leu Gly Leu Leu Val His
195          660          665
196

```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 712 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

RAW SEQUENCE LISTING PATENT APPLICATION US/08/240,410

DATE: 02/10/1999
TIME: 15:52:37

INPUT SET: S30581.raw

206
207 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:
208
209 CCGCCTGCCG CCCAGGGCCA CCTAAGGAGC GGATNCTANN TCTTCGCCGT CTGGAAAGGC 60
210 CATGTAGGGC AGGACCGGGT GGACTTTGGC CAGACTGAGC CGCACACGGT GCTTTTCCAC 120
211 GAGCCAGGCA GCTCCTCTGT GTGGGTGGGA GGACGTGGCA AGGTCTACCT CTTTGACTTC 180
212 CCCGAGGGCA AGAACGCATC TGTGCGCAG GTGAATATCG GCTCCACAAA GGGGTCCTGT 240
213 CTGGATAAGC GGGACTGCGA GAACTACATC ACTCTCCTGG AGAGGCGGAG TGAGGGGCTG 300
214 CTGGCCTGTG GCACCAACGC CCGGCACCCC AGCTGCTGGA ACCTGGTGAA TGCACTGTGG 360
215 TGCCACCTTG GCGAGAGTGG AGGCTACGCC CCCTTCAGCC CGGACGAGAA CGTCCCCTGG 420
216 TTCTGTTTTG AAGGGGACGA AGTGTATTCC ACCATCCGGA AAGCAAGGAA TTACAATTGG 480
217 GAAGATCCTC GGTTCGCGCG CATCCGGGGC GAGAGTGAGC TGTACACCAG TGATACTGTC 540
218 ATGCAGAACC CACAGTTCAT CAAAGCCACC ATCGTGCACC AAGACCAGGC TTACGATGAC 600
219 AAGATCTACT ACTTCTTCCG AGAGGACAAT CCTGACAAGA ATCCTGAGGC TCCTCTCAAT 660
220 GTGTCCCCTG TGGCCAGTT GTGCAGGGGG GACCAGGGTG GGGAAAGTTC AN 712
221

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 215 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

234 Gly Gln Asp Arg Val Asp Phe Gly Gln Thr Glu Pro His Thr Val Leu
235 1 5 10 15
236 Phe His Glu Pro Gly Ser Ser Ser Val Trp Val Gly Gly Arg Gly Lys
237 20 25 30
238 Val Tyr Leu Phe Asp Phe Pro Glu Gly Lys Asn Ala Ser Val Arg Thr
239 35 40 45
240 Val Asn Ile Gly Ser Thr Lys Gly Ser Cys Leu Asp Lys Arg Asp Cys
241 50 55 60
242 Glu Asn Tyr Ile Thr Leu Leu Glu Arg Arg Ser Glu Gly Leu Leu Ala
243 65 70 75 80
244 Cys Gly Thr Asn Ala Arg His Pro Ser Cys Trp Asn Leu Val Asn Ala
245 85 90 95
246 Leu Trp Cys His Leu Gly Glu Ser Gly Gly Tyr Ala Pro Phe Ser Pro
247 100 105 110
248 Asp Glu Asn Val Pro Trp Phe Cys Phe Glu Gly Asp Glu Val Tyr Ser
249 115 120 125
250 Thr Ile Arg Lys Ala Arg Asn Tyr Asn Trp Glu Asp Pro Arg Phe Arg
251 130 135 140
252 Arg Ile Arg Gly Glu Ser Glu Leu Tyr Thr Ser Asp Thr Val Met Gln
253 145 150 155 160
254 Asn Pro Gln Phe Ile Lys Ala Thr Ile Val His Gln Asp Gln Ala Tyr
255 165 170 175
256 Asp Asp Lys Ile Tyr Tyr Phe Phe Arg Glu Asp Asn Pro Asp Lys Asn
257 180 185 190
258 Pro Glu Ala Pro Leu Asn Val Ser Arg Val Ala Gln Leu Cys Arg Gly

PAGE: 1

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION US/08/240,410

DATE: 02/10/1999
TIME: 15:52:37

INPUT SET: S30581.raw

Line	Error	Original Text
28	Wrong application Serial Number	(A) APPLICATION NUMBER: TO BE ASSIGNED
30	Wrong Classification	(C) CLASSIFICATION: UNKNOWN